

BIOPS Interactive

An e-Learning Platform Focused on Protein Structure and DNA

Enrico Pontelli^{1*}, Jorge Pinto¹, Xiaoxiao Qin¹, Jing He¹, David Bevan²,
Norah MacCuish³, John MacCuish³, Mitch Chapman³, David Moreland⁴

¹ Dept. Computer Science, New Mexico State University, Las Cruces, NM 88003, USA

² Dept. Biochemistry, Virginia Tech University, Blacksburg, VA 24061, USA

³ Mesa Analytics & Computing, INC., Santa Fe, NM 87501

⁴ IDSC Biotech Network, Chelsea, MI 48118

Email: epontell@cs.nmsu.edu

*Corresponding author

Abstract: One of the difficulties in teaching basic molecular biology concepts to the students with little biological background is the lack of hands-on exercises that combines the challenges of the concepts with visualization and immediate feedback. BIOPS Interactive is a web-based interactive learning environment for molecular biology that complements traditional lecture and text coursework. It includes a set of exercises of basic concepts about molecules through the use of interactive graphics and scientific visualization. Students can gain immediate feedback of their exercises, and the instructor can get the real-time feedback about students' performance. In addition to the existing exercises, BIOPS provides a simple platform for instructors to create exercises of their own, or to modify existing ones. The preliminary feedback from BIOPS users has been positive.

Keywords: molecular biology, web-based education, interactive animation.

Introduction

The need for interdisciplinary education with respect to the field of biology has grown significantly during the past decade, fueled by the growing popularity of biologically motivated courses in disciplines like mathematical and computer science, and the prominent role that bioinformatics and computational methods have gained in biological investigations (e.g., (Johnson, 2001)).

The introduction of basic concepts from molecular biology to students with limited biology background (e.g., beginning biology undergraduate majors and students from other disciplines) is a challenging problem. Molecular biology concepts such as protein and DNA are often blended with the basic chemistry, 3-dimensional geometry, and biological relevance. Because of the underlying complexity of these concepts, it is vital to go beyond the text and pictures that can be found in a textbook.

The literature has also highlighted the importance of interaction and constructivism in the teaching of traditional sciences (Driver *et al.*, 1994; Yeany, 1991). In particular, several studies have underlined the need to introduce components of creative enquiry and collaborative learning in

introductory biology courses (Allen and Tanner, 2005; Lord, 1997; Millen, 2003). In presenting fundamental biology concepts to a novice student audience, it is desirable to have a set of well-designed exercises to challenge the students using an *interactive* and *graphical* environment. The literature has highlighted the importance of providing continuous feedback on the progress in the learning process, and the great impact that interactive environments have on retaining students' attention (Windelspecht, 2001). A natural consequence of this is the need for the instructor to be able to easily modify the exercises to address the needs of the class.

Computational tools offer solutions to enable viewing and investigating molecular biological entities (Honts, 2003; Musante, 2004). These tools are interesting, as they combine 3-dimensional visual representations with interactive behavior – e.g., selecting and deselecting parts of a molecule, zooming and rotation. Nevertheless, many viewers are often awkward to use and not designed to serve the educational mission – being designed as independent tools to be used by expert scientists.

This paper presents a novel tool to address these problems, called *BIOPS Interactive* (or, simply, *BIOPS*). The tool proposes to narrow the gap between theory and practice, students and instructors

. BIOPS is designed to enhance the learning of fundamental concepts from molecular biology, and it is meant to be used in introductory courses for biology majors as well as students for students that are approaching the field of biology from other disciplines (e.g., Computer Science students).

BIOPS Interactive is a web-based learning environment that enables students to practices with concepts in molecular biology – with particular focus on DNA and protein structure. BIOPS complements traditional classroom learning, by offering opportunities for both self-paced and synchronous lab

activities. Through the use of scientific visualizations the students are engaged in the discovery of biology concepts, by *interacting* with the problems. BIOPS is also designed to provide in-class interaction functionalities, allowing instructors to post problems to the class and interactively observe the class progress in the problem resolution. Instructors can create their own material and exercises, or modify existing BIOPS modules. BIOPS is a web-based (Figure 1) learning environment, where students can go online to exercise learned theory in a new interactive way.

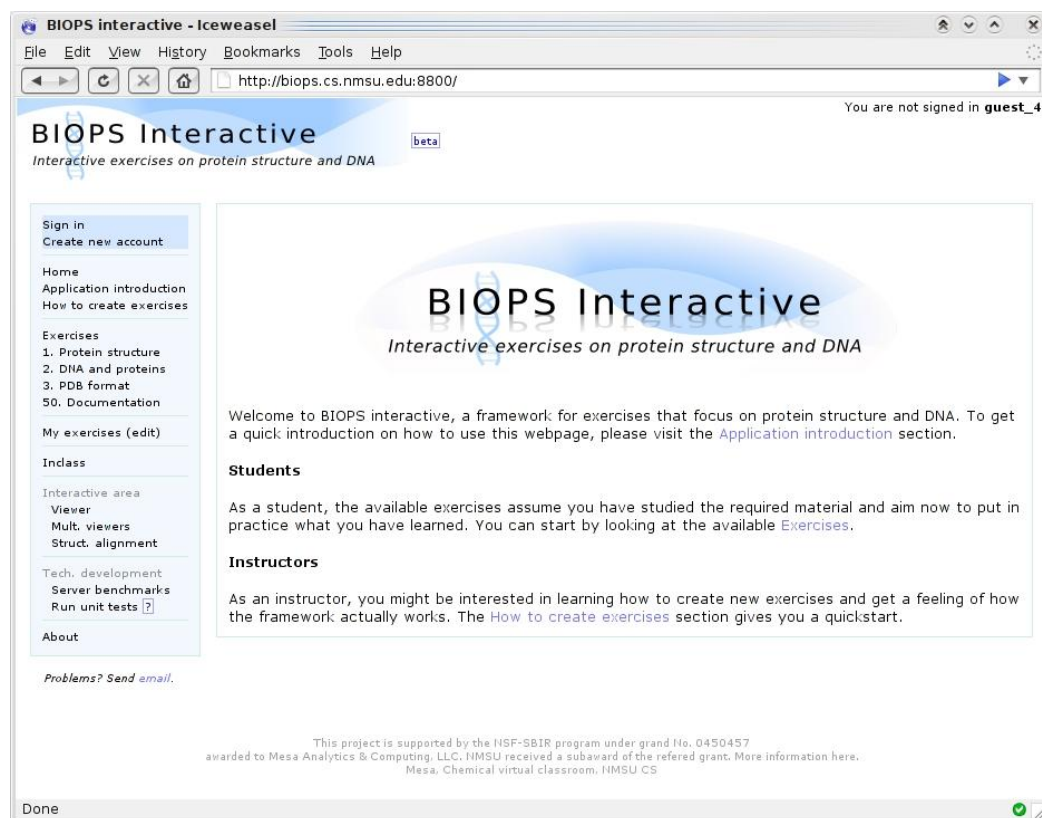


Figure 1: BIOPS website front page.

In BIOPS, the interactivity is achieved by proposing to the students exercises and online tools with direct visualization and interaction, and enabling hands-on practice of the content acquired in traditional textbook-based lectures. Software tools and a dedicated e-learning platform of exercises, with molecular biology content, have been developed, and these are discussed in the rest of this manuscript.

Materials and Methods

Basic concepts in molecular biology, like amino acid side-chain properties and protein secondary and tertiary structure, can be more easily acquired by biology and non-biology students if such concepts

can be related to a visual presentation. Traditional book-based learning relies on the use of static images to illustrate key concepts. Static images have the drawback of offering a single and flat presentation of concepts that are inherently three-dimensional.

The domain of computational biology offers a variety of resources, such as data repositories, search tools, and visualization tools; these resources have the potential to be useful in the educational setting, offering dynamic, mechanical, and operational views of biological entities. Nevertheless, the majority of such resources have not been designed for educational purposes, and their use in a classroom setting requires integration and refactoring.

A first step in addressing this problem is represented by the adoption of software tools that generate a three-dimensional presentation of molecules, with the ability to dynamically interact with the presentation, e.g., by performing rotation, translation, and zooming. In addition, the software tools are expected to link the graphical representation to other representations of the same concepts (e.g., nucleotide and amino acid sequences, secondary structure components) and place them in the context of a learning process.

Sequence and protein data (including information about secondary structures) are readily available in several repositories (e.g., the Protein Data Bank (PDB) (Berman *et al.*, 2000)). The data in these repositories has a fairly complex representation, it has a purely textual representation, and it is not readily usable as an educational instrument. A popular tool that can be used to provide graphical representation is *Jmol* (Herraez, 2007), a Java application that allows a user to interact with molecular models through three-dimensional motion graphics.

In BIOPS, *Jmol* is used as the visualization component, but its capabilities to provide interaction have been enhanced by integrating it in a coherent framework, and enhancing its capabilities to enable interaction with other tools and resources (such as PDB) – via control modules, programmed using a web scripting language. The interactivity helps the user, for instance, to select parts of a protein, highlight a chain, or visualize the amino acids that compose a helix.

Figure 2 shows an example of the integrated display of sequence and structure information for the student. This design can link the symbolic notation of amino acids or nucleic acids on the sequence to their 3-dimensional location and physical properties shown in the molecular viewer. This type of connection is very helpful for the students with little biology background to quickly realize the concepts behind the symbols. In this DNA-protein complex – named 1SKN in the PDB and representing a DNA-binding

domain in eukaryotic transcription factors that specifies mesoderm in *C. elegans* (Rupert *et al.*, 1998) – two nucleotides of chain A are highlighted in the viewer (in yellow) by a user, as a result of making the selection of the “CC” symbol in the sequence module (in red, bottom). Both the sequence and the viewer display identical colors on the secondary structure for easy identification. Shortcut buttons for movement actions are immediately below the viewer to aid the user.

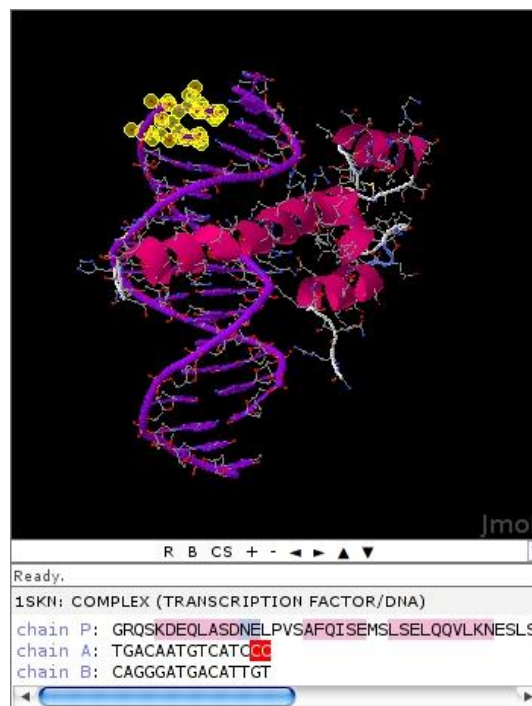
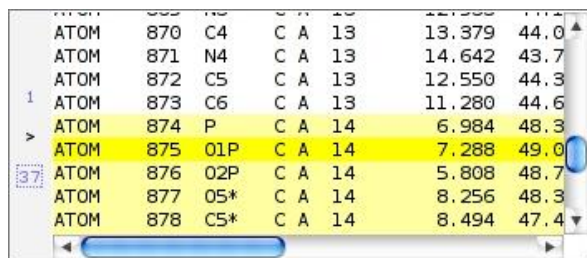


Figure 2: Sequence display module, showing the structure sequence displayed in *Jmol*

The tool is directly linked to the PDB repository, from which the data can be directly downloaded and used. In order to facilitate the interaction with PDB, BIOPS includes a module with which students can become familiar with the PDB format (Figure 3). The module is connected to the sequence and the viewer modules.



ATOM	870	C4	C	A	13	13.379	44.0
ATOM	871	N4	C	A	13	14.642	43.7
ATOM	872	C5	C	A	13	12.550	44.3
ATOM	873	C6	C	A	13	11.280	44.6
ATOM	874	P	C	A	14	6.984	48.3
ATOM	875	O1P	C	A	14	7.288	49.0
ATOM	876	O2P	C	A	14	5.808	48.7
ATOM	877	O5*	C	A	14	8.256	48.3
ATOM	878	C5*	C	A	14	8.494	47.4

Figure 3: PDB file format module, highlighting the selected atoms in Jmol

Results

BIOPS for Students and Instructors

From a student perspective, BIOPS aims at presenting exercises with a rich interactive interface that motivates the student to further understand and explore the topic. The available exercises are presented to the students in a simple menu, organized in separate sections based on topics.

The BIOPS platform supports different degrees of interaction between the student and the instructor (i.e., the creator of the exercise). The students can provide solutions to each exercise as well as provide feedback about the content of the exercise – in the form of textual comments and a simple rating of 1 to 5 – to help instructors improve the content. In turn, the instructor can keep track of progress made by the class and by the individual students in solving the exercises.

Instructors can use BIOPS to create exercises with interactive content. These exercises can be deployed on the web, for self-paced student practice,

or used directly in the classroom. An instructor can receive students answers in real-time during the class, in order to monitor progress and to help to decide if the class understood the material, or if a review should be made.

BIOPS is available online. It includes a direct viewer section where proteins and DNA can be accessed through their PDB identifier, and a multi structural alignment tool accessible through the web. In the successive sections we will illustrate the features of BIOPS through examples of its use.

An Exercise Performed by the Student

Figure 4 shows an exercise that introduces amino acids. This exercise is designed to test if a student can distinguish the 20 different amino acids when their chemical structures are given. Through this exercise, a student is challenged with the concepts of side chains and their different properties. The student is asked to look at the viewer displaying a dipeptide in 3-D. One of the amino acids, arginine, is given, while selecting the other one is the student's task. There is a table of amino acids on the right, which allows the student to view a two-dimensional representation of one or more specific amino acids. By comparing the 3-D representation with the 2-D representations, the student can analyze different amino acids and select the one that matches the best. The ability to manipulate the 3-D representation, via rotation and zooming, allows the student to determine a view of the dipeptide that facilitates the comparison. This process engages the student in finding the answer using his/her own intuitive analytical skills.

1.1.1. Exercise

Amino acids are the building blocks of proteins. An amino acid has a central alpha carbon with an amino group, a carboxyl group, and a side chain (R-group) attached to it. The 20 standard amino acids have different properties determined by their side chains. The sequence of amino acids is the primary structure of a protein.

Figure 4: An exercise, with the amino acids table presented in 2-D to aid the identification of the 3-D representation of a molecule

After identifying the amino acid, the student is asked to select the answer from a list of possible answers (Figure 5).

Figure 5: Feedback returned after submitting the answer to an exercise

An Exercise Used by the Instructor In-Class

The web-based nature of BIOPS allows us to take advantage of internet-enabled classrooms, and it enables students to participate in guided in-class sessions. With this scenario in mind, the instructor initiates a new in-class session and chooses the exercise to be solved. Students connect to the instructor session and try to solve the problems. The instructor session automatically presents group and individual statistics regarding the answers submitted by the students connected to the session. If the instructor decides to change the session exercise, the session on each student's computer will automatically change as well, providing a synchronized focus on the lecture.

Figure 6 shows a table where students tried to access a session and tried to solve three sections of the exercise

section	correct	wrong	unanswered
section_1	2	0	0
section_2	0	1	1
section_3	1	0	1
total: 2 users	50%	17%	33%

user name	correct	wrong	unanswered	submissions
guest_31	2	1	0	3
guest_27	1	0	2	1

Figure 6. Statistics for an in-class exercise.

After the results have been submitted, the instructor can decide, based on the feedback returned, what material to cover next. The immediate feedback gives the instructor a way of dynamically and objectively adjusting the next step in the class (e.g., revise materials, propose new exercises on the same topic, move to a new topic).

The ideal use of this feature is in a computer lab setting, where students practice the material learned in class. Furthermore, BIOPS enable interactive remote participation of students to lab sessions, instead of the passive model promoted by traditional online management systems (e.g., Moodle).

Bringing the Community into the Loop: Exercise Creation

In recent years, the web has gained a prominent role as a framework the "grassroots" development of repositories of community-provided knowledge – this trend has been referred to with different terms, such as Web 2.0 (Shuen, 2008). Web 2.0 technologies offer the ability to share knowledge and provide community-based evaluation of data and hypotheses (e.g., CBioC (Baral, 2006)).

The potential for involvement of a community of educators in the development of a shared educational infrastructure has been recognized by several authors (Alexander, 2006; Vonderwell, 2008). BIOPS provides the ability for a community to contribute exercises, which can be used and shared with other instructors. Exercises are stored in BIOPS in a XML format, making it easy for educators to use existing exercises as templates and modify them as suited. The feedback mechanism enables continuous improvement of exercises and the creation of a community-based repository of user-ranked exercises.

To support this approach, BIOPS provides a repository of exercises, along with an online section where new exercises can be created. Any registered user can create exercises and save them for later

reference and share them with other BIOPS users. A relatively limited amount of computing knowledge is required to develop exercises, and online documentation is provided on how to create them in a simple manner, from simple questions to more advanced features that require scripting.

BIOPS provides a graphical user interface for the interactive development of new exercises. Each exercise is described using a specialized XML format specifically developed within BIOPS.

An exercise is composed of two parts. The first part describes the question to be proposed to the student. The question can be formatted using standard HTML; in addition, the problem description can include a script section, used to perform processing of the answers and to invoke Jmol. The second component of an exercise includes information concerning the correct answer to the problem. A new exercise can be created by modifying one of several templates provided by BIOPS.

We noticed that teaching the basic molecule structures often involves teaching the data that represents the structure. BIOPS encourages students to first understand what a molecule is and have a hands-on experience with the molecule through the 3-dimensional viewer, and only later learn about the underlying data. One of the exercises was designed to enable loading a protein structure, visualizing it, and manipulating it.

The interactive area allows also access to an online structural alignment tool, enabling the display of the alignment of up to four chains. As seen in Figure , after the alignment is performed, the user can perform the desired selection and visualize it in two viewers. One of the viewers (left) shows the overall alignment, the second (right) shows only the selection made.

challenge the understanding of the structural data provided in PDB format. Custom-made PDB files may be uploaded by instructors who want to create an exercise that shows molecules, ranging from simple molecules up to multiple proteins within the same viewer. More advanced displays, with two or more viewers for the same PDB file are also possible (Figure).

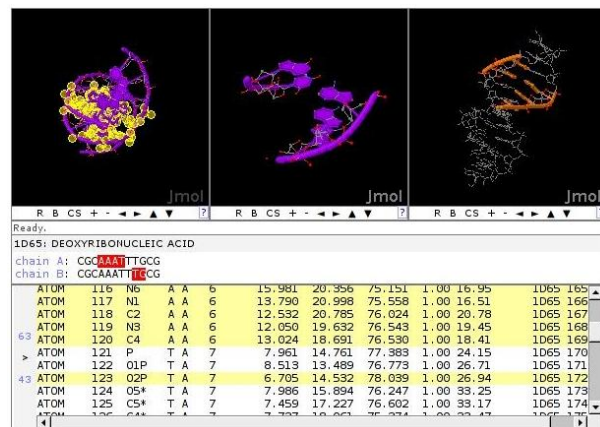


Figure 7: Multiple viewers showing the same molecule differently

Interactive Areas

BIOPS provides users with the ability to open interactive areas on the web site. A common use of the interactive area is to connect to PDB and to

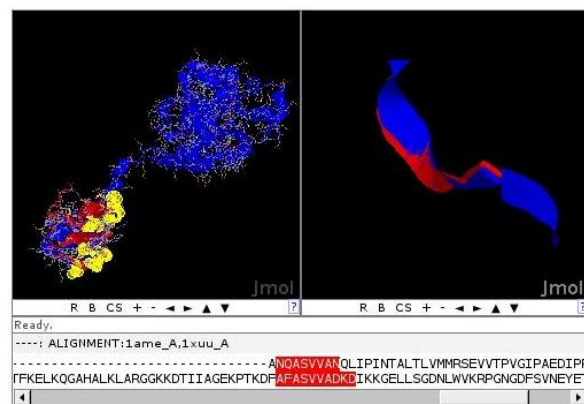


Figure 8: Interactive use of the alignment tool

System Organization and Implementation

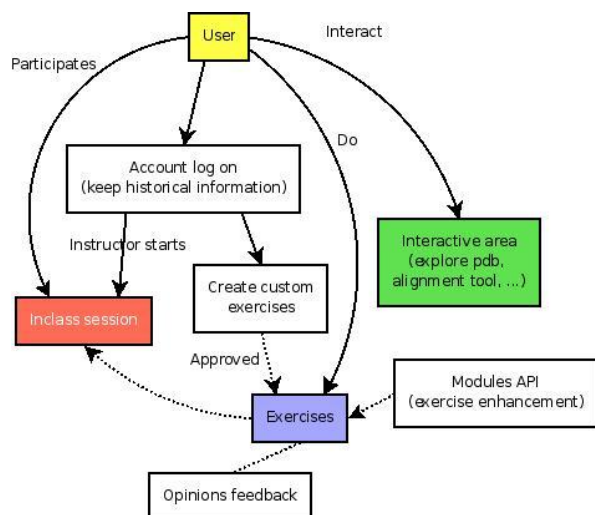


Figure 9: System design overview

The key components and interactions of the BIOPS framework are depicted in Figure 9. The role of the user in the system and the progress in problem solving and class participation are managed through a centralized accounting mechanism. An account can be freely created, and it allows BIOPS to maintain user profiles and user-specific historical information, such as answers submitted. The accounting mechanism for the instructor includes the ability of tracking exercises developed and historical data concerning success and failure rates per problem.

The core of the system is represented by the collections of modules used to manage exercises. Any user of the system can create exercises. Exercises are stored in a central repository, where they can be retrieved, reviewed, and included in organized collections for in-class sessions.

The exercises management component provides a graphical user interface and an XML language for the development of new exercises. The specialized language allows access to a collection of enhanced modules Application Programming Interface (API), which supplies the improved interactive resources mentioned earlier.

Figure 10 shows a more technical system overview of BIOPS. The web pages requested by users from the Internet are enhanced with Javascript and AJAX, making use of Java Applets, like Jmol, to show 3-D representation of molecules.

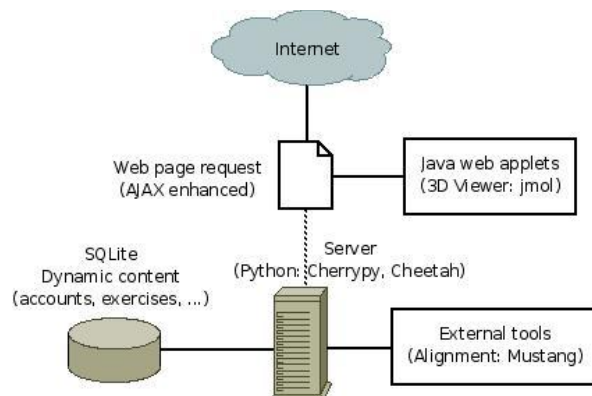


Figure 10: Technical system overview

The main BIOPS server has been developed in Python, using CherryPy as web server and Cheetah (Hellegouarch, 2007) as the template engine. The repositories of exercises and user accounts have been implemented using a lightweight SQL-based relational database. External tools are used in real time by the server to provide interactive functionalities.

Topic	Question number	Question	Score	Spring 08	Spring 09
Environment Used:	1	What type of computer did you use to access BIOPS?			
	2	What operating system is installed on your computer?			
	3	Which web browser did you use to access BIOPS?			
BIOPS Interface Design	4	Was the interface easy to use?	Very Easy Easy So-So Not Easy Difficult	4.8/5	4.5/5
	5	Did you know what you needed to do?	Very Obvious Obvious So-So Confusing Not Obvious	4.2/5	4.1/5
	6	Were you able to figure out how to manipulate structures in Jmol?	Very Easy Easy So-So Not Easy Difficult	4.1/5	4.0/5
	7	Were the Jmol widgets (e.g., amino acid structures) useful?	Very Useful Useful So-So Somewhat Useful Not Useful	4.9/5	4.6/5
BIOPS Content	8	Was the content appropriate to learn about protein structure?	Very Appropriate Appropriate So-So Inappropriate Very Inappropriate	4.4/5	4.6/5
	9	Did you gain better understanding of protein structure after working through the exercise?	Much Better Better About the Same	4.6/5	4.6/5
Clarity of Content Presentation	10	Were you able to answer the questions in the tutorial correctly and understand why you were correct on your first attempt?	Most of them Many Some A Few None	4.6/5	4.6/5
	11	Were you able to answer the questions in the tutorial correctly and understand why you were correct after multiple attempts?	Most of them Many Some A Few None	4.2/5	3.9/5
General Experience Evaluation	12	What did you learn by using BIOPS Interactive?			
	13	What parts of BIOPS Interactive did you like best?			
	14	What parts of BIOPS Interactive did you like least?			
	15	Do you have suggestions on how to improve the educational value of BIOPS Interactive?			

Table 1. Evaluation questionnaire and summary results.

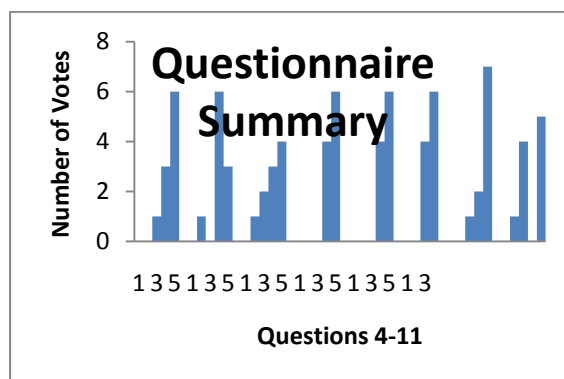
Preliminary Evaluation

The development of a first release of BIOPS has been recently completed and informal evaluations have been performed.

In July of 2007, BIOPS was presented at the 21st Symposium of The Protein Society in a workshop for educators. A total of 35 educators participated in the session. After the presentation, the workshop allowed hands-on activities, where participants interacted with exercises in an in-class session. The experiment was successful both technically and as a learning event, where the participants engaged in interacting with the exercises and followed the session. The feedback was positive with several expressions of interest for classroom use of BIOPS.

BIOPS was deployed at Virginia Tech, as part of two graduate level classes in Computational Biochemistry (Spring 2008 and 2009). The background of the students was primarily Computer Science. We distributed to the students an evaluation form with 15 questions. The questionnaire is summarized in Table 1. The students had never seen BIOPS before. The evaluation included sections requesting an evaluation of the interface design, including usability of the modules. The evaluation also covered the perceived strengths and weaknesses of BIOPS, especially concerning the effectiveness as a learning instrument.

The evaluation indicated a high level of satisfaction in the use of the system, including praise for the easy to use interface and the easy way to solve and develop exercises. References were made, in particular, to the BIOPS modules where sequences can be selected and highlighted. Regarding the strengths of BIOPS, the ability to interact with the molecules and with the three-dimensional representation was often cited. The evaluation overwhelmingly confirmed that BIOPS helped the students understanding the material – students stressed the impact of being able to associate theoretical concepts to the graphical structures (e.g., representations of amino acid structures or by highlighting secondary structures). The main weakness reported was the need to expand the exercise repository – something that will naturally occur as the BIOPS user base expands. Figure 12 summarizes the scores given to questions 4-11 for the Spring 2009 edition of the class (10 students). Analogous results were collected in Spring 2008 (15 students). The scores of each column corresponds to one of the categories (1=worst, 5=best).



Acknowledgments

This project has been supported by NSF grants 0450457 and 0420407. The authors wish to thank the anonymous reviewers for their comments.

References

ALEXANDER, B. 2006. Web 2.0: A New Wave of Innovation for Teaching and Learning. *EDUCAUSE Review*, 33-44.

ALLEN, D. AND TANNER, K. 2005. Infusing Active Learning into the Large-enrollment Biology Class: Seven Strategies, from the Simple to the Complex. *Cell Biology Education* 4, 262-268.

BARAL, C. 2006. CBioC: Annotate While You Read. *Science* 312, 1721.

BERMAN, H.M., WESTBROOK, J., FENG, Z., GILLILAND, G., BHAT, T.N., et al. 2000. The Protein Data Bank. *Nucleic Acid Research* 28, 235-242.

DRIVER, R., ASOKO, H., LEACH, J., MORTIMER, E. AND SCOTT, P. 1994. Constructing Scientific Knowledge in the Classroom. *Educational Researcher* 23, 5-12.

HELLEGOUARCH, S. 2007. *CherryPy Essentials: Rapid Python Web Application Development*. Packt Publishing.

HERRAEZ, A. 2007. *How to Use Jmol to Study and Present Molecular Structures*. Lulu.com.

HONTS, J.E. 2003. Evolving Strategies for the Incorporation of Bioinformatics within the Undergraduate Cell Biology Curriculum. *Cell Biology Education* 2, 233-247.

JOHNSON, G. 2001. All Science is Computer Science. *The New York Times*.

LORD, T.R. 1997. A Comparison Between Traditional and Constructivist Teaching in College Biology. *Innovative Higher Education* 21, 197-216.

MILLEN, J.W. 2003. Factors Affecting the Implementation and Use of Technology in Teaching Biology Courses in Florida's Community Colleges. Curtin University of Technology.

MUSANTE, S. 2004. Using Bioinformatics in the Undergraduate Classroom. *American Institute of Biological Sciences - Eye on Education*.

RUPERT, P.B., DAUGHDRILL, G.W., BOWERMAN, B. AND MATTHEWS, B.W. 1998. A New DNA-binding Motif in the Skn-1 Binding Domain-DNA Complex. *Nat. Struct. Biol.* 5, 484-491.

SHUEN, A. 2008. *Web 2.0: A Strategy Guide*. O'Reilly Media, Inc.

VONDERWELL, S. 2008. Web 2.0 and Implications on Higher Education. *Annual Meeting of the MWERA*. Columbus, OH.

WINDELSPECHT, M. 2001. Technology in the Freshman Biology Classroom: Breaking the Dual Learning Curve. *The American Biology Teacher* 63, 96-101.

YEANY, R.H. 1991. A Unifying Theme in Science Education? *NARST News* 33, 1-3.